

Felügyelt tanulás kis mintaszám alapján történő gesztusfelismeréshez spiking neurális hálókkal

Gyöngyössy Natabara, Domonkos Márk, Dr. Botzheim János, Dr. Korondi Péter

Absztrakt

Ez a poszter egy új, felügyelt tanulási algoritmust mutat be spiking neurális hálózatokhoz. Az algoritmus ötvözi a Hebb-tanulást és a legkisebb négyzetek módszerét (LMS), és jól működik kis tanító adatkészletek és kis ismétlésszám esetén. A javasolt módszert az ember-robot interakcióban alkalmazzuk Kodály Zoltán munkája alapján, a szolmizációs jelek megkülönböztetésében. Az MNIST adatbázist referenciatesztként is használjuk annak ellenőrzésére, hogy a javasolt algoritmus képes-e felülmúlni a sekély ANN architektúrákat. A robottal végzett kísérletek szintén ígéretes eredményeket hoztak az emberi kézjelek helyes felismerésében.

Bevezetés

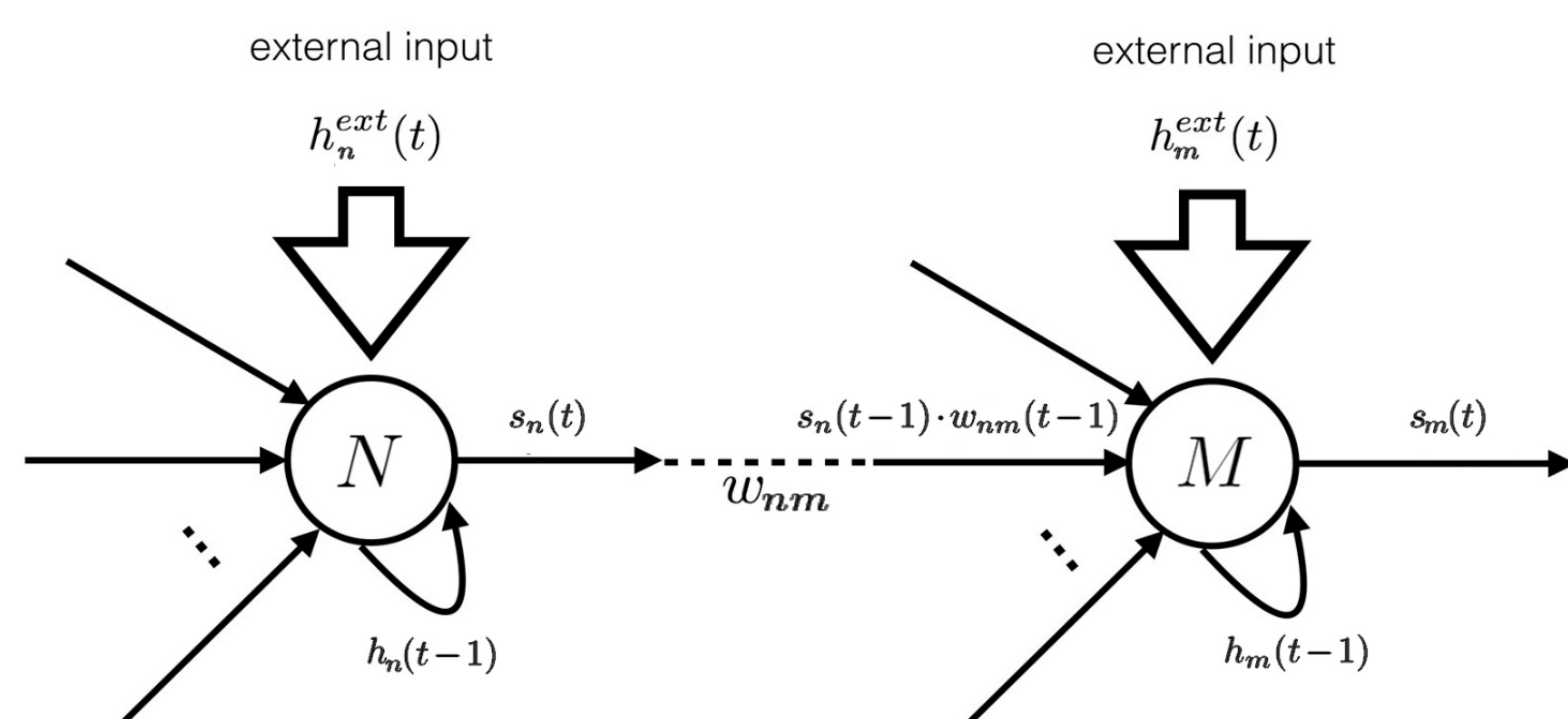
Korunk ipari folyamatai során közös ember-robot munkatér van kialakulóban, melyhez a merev irányítási módszereket finomabb, rugalmas modellekkel kell helyettesítenünk. Ezt a témát járja körül a kognitív robotika területe is, mely magában foglalja az ember-robot interakciók felépítését is. Jelen kutatásban egy spiking neurális hálózat valósítja meg a gesztusfelismerés feladatát, a robot vezérlése érdekében. A vezérlőjelek a Kodály Zoltán által bevezetett szolmizációs kézjelek [1], melyek alapján a robot egy zongora billentyűit üti le az operátor mozgatait lekövetve.

Módszerek

A spiking neurális hálózatok alapvetően eseményszinten modellezik az emberi neuronokat [2]. Így a membránpotenciál töltődését, a neuron tüzelését, illetve gátolt állapotát szimulálják. A tanulás elsősorban a Hebb-féle posztulátumok alapján működő STDP (Spike-timing dependent plasticity) algoritmus alapján történik ilyen hálózatok esetén, mely felügyelet nélkül valósít meg tanulást.

Ez a módszer a legkisebb négyzetek módszerével (LMS) [3] ötvözve felügyeltté tehető olyan módon, hogy a kimeneti neuron tüzelési idejének elvárttól való eltérése alapján definiálunk hibát. A bemenetet a bejövő tüzelésekből összegezve (integrálva) állítjuk elő, amennyiben ezek hiányoznak, az STDP algoritmus negatív ága [4] aktiválódik, így mind az LMS, mind az STDP algoritmusnál pontosabb lesz ez a módszer.

A robot irányításánál a kezlet szín alapú szegmentációval, illetve a háttér kivonásával különítjük el környezetétől. A jobb kéz a szolmizációs jelet mutatja, ezt egy 200 pixel élhosszúságú négyzet alakú szürkeárnyaltos kép reprezentálja, míg a ritmusjelet a bal kéz szolgáltatja a szegmentáció után visszkapott objektum magassága alapján döntve a lenyomásról.



1. ábra: Spiking neuronok kapcsolata (egyszerűsített non-leaky integrate & fire modell)

Eredmények

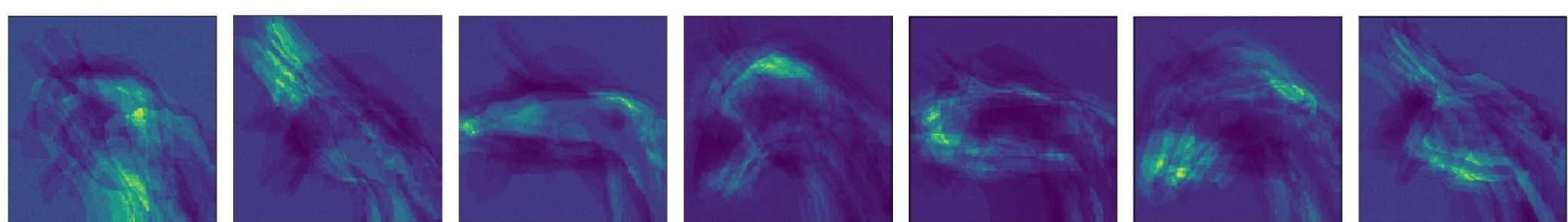
Törekvéseink közé tartozott, hogy a gesztusfelismerést minél kisebb, sekélyebb, alacsony erőforrásigényű hálózattal valósítsuk meg. Így egy két rétegű hálózat került tanításra, melynek első rétegére bemenetként a vizsgált kép szürkeárnyaltos pixelfényességei kerülnek. A detektált kategóriát a kimeneti réteg először tüzelő neuronjának sorszáma jelenti.

Vizsgálataink alapján az MNIST [5], kézzel írott számjegyekből álló adatbázis tanító mintáinak 5%-át felhasználva a teljes tesztsonon 3 tanítóciklus után 81%-ban helyesen kategorizált hálózatunk, amennyiben STDP-LMS tanítást alkalmaztunk. Összehasonlításképpen, egy 3-rétegű klasszikus neurális hálózat (multi-layer perceptron, MLP) pontossága 77% volt ugyanilyen paraméterek mellett, míg a sima STDP algoritmus felismerési aránya 60% alatt mozgott.

A szolmizációs kézjelek felismerése esetén (kategóriánként 100 minta alapján) hálózatunk 93%-os konfidencia mellett működött ugyanazon felhasználó esetén. Amennyiben a felhasználó számára visszacsatolást biztosítottunk, úgy az emberi alkalmazkodóképesség kihasználásával ez a teljesítmény 98%-osra nőtt. Ezáltal az ember-robot kölcsönös tanulása (kétirányú adaptáció) is mérhetően javította a feldolgozás sikerességét. A hálózat teljes újratanítása gyors, 5 percet vesz igénybe az operátor betanításával együtt, továbbá másodpercenként minimum 20 képkocka feldolgozására képes.



2. ábra: A zongorázó robot (Bal kéz a leütések idejét, jobb kéz a leütött billentyűt határozza meg)



3. ábra: A két réteg közötti súlyokban reprezentált információ az egyes kézjelek esetén

Konklúzió

Spiking neurális hálózatokkal megvalósítható a klasszikus (MLP) hálózatoknál kevesebb tanulómintából is konvergáló sekély (2 rétegű) hálózat. Ez a hálózat tanítható az STDP és az LMS módszer ötvözetével, az STDP-LMS algoritmussal. A hálózat hiperparaméterei érzékenyek, további optimalizációs eljárások szükségesek lehetnek kiválasztásukkor. Személyre szabható, gyors betanítású neurális hálózatok könnyedén létrehozhatók ember-robot interakcióhoz, mely során a kölcsönös tanulás szerepe sem elhanyagolható.

Hivatkozások

- [1] E. Hegyi, Z. Kodály, „Solfege According to the Kodály-concept: Chapters I. to V.”, Zoltán Kodály Pedagogical Institute of Music, 1975, vol. 1.
- [2] W. Gerstner, W. M. Kistler, „Spiking neuron models”, Shaftesbury Road, Cambridge: Cambridge University Press, 2002.
- [3] B. Widrow, M. E. Hoff, „Adaptive switching circuits”, IRE WESCON Conv. Record, vol. 4, pp. 96–104, 1960.
- [4] S. R. Kheradpisheh et. al., „STDP-based spiking deep convolutional neural networks for object recognition”, Neural Networks, vol. 99, pp. 56–67, 2018.
- [5] Y. LeCun, C. Cortes, „MNIST handwritten digit database”, <http://yann.lecun.com/exdb/mnist/>, 2010. [Online]